REPORT DOCUMENTATION PAGE

Form Approved OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
January 1999

3. REPORT TYPE AND DATES COVERED
1998 Annual Report

4. TITLE AND SUBTITLE
1998 Annual Report, Ecological Modeling and Simulation Using Error and Uncertainty Analysis Methods (Project CS-1097)

5. FUNDING NUMBERS
N/A

6. AUTHOR(S)
A.W. King, T.L. Ashwood, B.L. Jackson, H.I. Jager, C. Hunsacker

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Oak Ridge National Laboratory

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
SERDP
901 North Stuart St. Suite 303
Arlington, VA 22203

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
N/A

11. SUPPLEMENTARY NOTES
The United States Government has a royalty-free license throughout the world in all copyrightable material contained herein. All other rights are reserved by the copyright owner.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release: distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT *(Maximum 200 Words)*

The principle objective of this project is to identify and implement moethods for the analysis of error and uncertainty of spatial data in spatially explicit ecological models. The development of these methods for spatial error and uncertainty analysis is done in coordination with Dr. George Gertner and SERDP Conservation Thrust Area Project SC-1096. Methods and tools developped and implemented by our project will be transferred to Project CS-1096 for incorporation within that project's error budget framework and toolbox. Our software tools will also be compatible with the specifications and requirements of DoD and SERDP land management tools and software like the U.S. Army Corps of Engineers' Land Management System (LMS) or others.

We have completed the first phase of a survey of existing methods for the analysis of error and uncertainty analysis in spatial data. From this survey we have formulated a general framework for spatial erro and uncertainty analysis of ecological models, and we have implemented software tools for treatment of error and uncertainty in categorical spatial data. We have begun the process of testing and refining these software tools.

14. SUBJECT TERMS
ecological modeling; error analysis; SERDP; SERDP Collection

15. NUMBER OF PAGES
21

16. PRICE CODE N/A

17. SECURITY CLASSIFICATION OF REPORT
unclass.

18. SECURITY CLASSIFICATION OF THIS PAGE
unclass.

19. SECURITY CLASSIFICATION OF ABSTRACT
unclass.

20. LIMITATION OF ABSTRACT
UL

1998 ANNUAL REPORT
Ecological Modeling and Simulation Using Error and Uncertainty Analysis Methods (Project CS-1097)

A. W. King
T. L. Ashwood
B. L. Jackson
H. I. Jager
Oak Ridge National Laboratory

C. Hunsaker
USDA Forest Service
Pacific Southwest Forest & Range Experiment Station

January 1999

19990525 009

DTIC QUALITY INSPECTED 1

Contents

List of Figures

1 Executive Summary

The Strategic Environmental Research and Development Program (SERDP) of the Department of Defense (DoD) is sponsoring research on the application of error and uncertainty analysis to ecological models used in military land-use management and decision support. Increasingly, the ecological models used in DoD ecosystem and land-use management are spatially explicit, relying on spatially distributed and georeferenced data as model input. The spatial inputs to these models, commonly in the form of geographical information system (GIS) data layers, have varying degrees of uncertainty associated with them. This uncertainty needs to be propagated throughout the entire modeling and simulation process so that: (1) model results can be presented to the ecosystem/land manager as a probability distribution of possible outcomes, and (2) the contribution of uncertainty in spatial data to overall model uncertainty can be quantified as part of an error budget analysis.

The principle objective of this project, SERDP Conservation Thrust Area Project CS-1097, is to identify and implement methods for the analysis of error and uncertainty of spatial data in spatially explicit ecological models. The development of these methods for spatial error and uncertainty analysis is done in coordination with Dr. George Gertner and SERDP Conservation Thrust Area Project CS-1096. Methods and tools (e.g., computer codes) developed and implemented by our project will be transferred to Project CS-1096 for incorporation within that project's error budget framework and toolbox. Our software tools will also be compatible with the specifications and requirements of DoD and SERDP land management tools and software like the U.S. Army Corps of Engineers' (USACE) Land Management System (LMS) or others.

We have completed the first phase of a survey of existing methods for the analysis of error and uncertainty analysis in spatial data. From this survey we have formulated a general framework for spatial error and uncertainty analysis of ecological models, and we have implemented software tools for treatment of error and uncertainty in categorical spatial data. We have begun the process of testing and refining these software tools.

We have selected Fort Hood, Texas as our case-study site. We will apply a spatially-structured avian population model to populations of black-capped vireo and golden-cheeked warbler at Fort Hood, and we will complete an error and uncertainty analysis of the model focusing on the contribution of error and uncertainty in spatial data. The case study will familiarize us with spatial data at Fort Hood, facilitate coordination with the needs and requirements of the Army Training and Testing Area Carrying Capacity (ATTACC) methodology and of CS-1096, and will, as a side benefit, contribute to the investigation of black-capped vireo and golden-cheeked warbler at Fort Hood. Our methods and tools (e.g., software) tested with the population model will be transferred to Dr. George Gertner and CS-1096 as part of Dr. Gertner's error budget approach for application to ATTACC at Fort Hood. Although ATTACC is not spatially explicit at present, the methods and tools developed from our population model case study will be ready for future versions of ATTACC that are spatially explicit and require the methods we are developing. In addition, separating the case-study development of our methods and tools from the application of these tools to ATTACC will provide for a more powerful independent test and validation of our methods, one which will facilitate transfer and incorporation into LMS.

2 Problem Statement and Background

Increasingly, ecological models used in ecosystem and land-use management are spatially explicit, relying on spatially distributed and georeferenced data as model input. The spatial inputs to these models, commonly in the form of geographical information system (GIS) data layers, have varying degrees of uncertainty associated with them. This uncertainty needs to be propagated throughout the entire modeling and simulation process so that: (1) model results can be presented to the ecosystem/land manager as a probability distribution of possible outcomes, and (2) the contribution of uncertainty in spatial data to overall model uncertainty can be quantified as part of an error budget analysis. The latter provides for cost effective allocation of resources to reduce uncertainty in model output and to minimize, to the extent possible, the range of potential outcomes the manager must evaluate.

Geographers and geostatisticians have been interested in the issue of uncertainty in spatial information for a long time (Burrough, 1986; Isaaks and Srivastava, 1989), and it remains a priority issue within those communities. Uncertainty in spatial data is, for example, one of the ten research priorities of the University Consortium for Geographic Information Science (UCGIS, 1996). Nevertheless, there are few readily available techniques and tools to address uncertainty in spatial data. Existing methods are largely found within the geostatistics and geographical informations system research communities, and these have been slow to penetrate into the ecological modeling community. Thus there is a particular need for research and development of methods and tools for spatial error and uncertainty analysis of ecological models.

Error and uncertainty in spatial data arise from a variety of sources, including natural variability, inaccuracies in geographic coordinates, measurement error and misclassification in the field, and errors that arise in the processing and interpretation of data (e.g., land-cover classification of remote sensing imagery). These and other sources of spatial data error have been well defined by Burrough (1986), Goodchild and Gopal (1989), and others. Operationally, from the perspective of developing and implementing general methods of spatial uncertainty analysis, it is useful to characterize spatial error and uncertainty as: (1) error in categorical data (e.g., soil or vegetation type) and (2) error in continuous quantitative data (e.g., vegetation height or population density). The most appropriate methods of spatial error analysis are fundamentally determined by these broad categories. General methods appropriate to these basic data types may then be refined if necessary to reflect differences in the source of that error or uncertainty.

We approach the problem of error and uncertainty in spatial data from the perspective of ecological modelers involved in applying ecological models to problems of ecosystem and land-use management. Thus our primary focus is on how to quantify error and uncertainty in spatial data (e.g., a GIS data layer) presented as input for an ecological model, how to propagate that uncertainty throughout the simulation process, and how to relate the resulting variability in model output to uncertainty in model input (e.g. which input is most responsible for variability in model output). Findings from the latter can be used to target methods of minimizing error and uncertainty in spatial data collection and processing to those spatial data with the greatest contribution to error and uncertainty in model output.

Beyond simply identifying available and appropriate methods, it is important that these methods be implemented as usable and practical tools. Methods and approaches must be translated into software and incorporated into modeling and analysis systems. The software developed as part of research and development in a 6.1 Basis Research project such as the one described here need not be, and are unlikely to be, the "polished" product distributed as part of a modeling or decision support system. Nevertheless, the selected approaches and methods of analysis should be implemented as functional tools, tested on and applicable to real world situations, and consistent with the design specifications of the systems for which they are ultimately destined.

2.1 Project Objectives

The principle objectives of this research project are to:

1. Identify, evaluate, and implement methods for quantifying error and uncertainty in spatial data used in ecological models.
2. Incorporate error and uncertainty in spatial data into a Monte Carlo framework for uncertainty and error analysis of spatially explicit ecological models.
3. Test and demonstrate the analytical framework, methods and tools with one or more case studies.
4. Transfer our methods and tools to Dr. George Gertner and CS-1096 for incorporation into the CS-1096 error budget framework and toolbox.
5. Make our methods and tools compatible with, and available for incorporation into DoD and SERDP land management tools and software, e.g., the Land Management System (LMS).

In this Annual Report we document our achievements and findings for the first year of the project (1998). As we reported at the 1998 In-Progress Review for the the Conservation Technology Thrust Area (May 20, 1999, Arlington, VA), our project is focussed on error and uncertainty analysis of spatial data as used in ecological models. This focus "bores in" on one component of Dr. George Gertner's error budget (re CS-1096) and, consequently, supports and complements that broader approach. Our focus on spatial error and uncertainty and ecological models has been coordinated with Dr. Gertner and CS-1096, and we will continue that coordination as we proceed.

3 First Year Objectives

Our milestones for 1998 were:

1. Define all potential errors to be considered
2. Initiate analytical/uncertainty methodologies to quantify the errors
3. Develop required model to quantify errors
4. Select monitoring-modeling system for case study
5. Evaluate and assess methodology (Go/no-Go decision)
6. Annual Interim Report to SERDP

In achieving these milestones our goals for the first year of the project were to:

1. Survey existing methods and tools of spatial error and uncertainty analysis
2. From this survey, identify and evaluate the most appropriate and efficient methods for quantifying error and uncertainty in spatial data used in ecological models.
3. Develop a general framework for spatial error and uncertainty analysis of ecological models
4. Begin the implementation of appropriate methods as software tools.
5. Select a case-study site and model (or models) with which to test and develop methods and tools.

4 Achievements

Funding was received at Oak Ridge National Laboratory (ORNL) on April 7, 1998. This report covers activities and accomplishments of the project for the period April 1998 through December 1998.

We have completed the first phase of our survey for existing methods and approaches for quantifying error and uncertainty in spatial data. This survey included the findings of September 1997 workshop at the National Center for Ecological Analysis and Synthesis (NCEAS) with the goal of synthesizing knowledge and analysis techniques for uncertainty in spatial data. Dr. Carolyn Hunsaker has headed this portion of our survey. An annotated outline of the proceedings of the NCEAS workshop is presented in Appendix A.

From our survey of existing methods, we have identified *stochastic simulation* as the most broadly applicable approach to incorporating spatial error and uncertainty into ecological models. Stochastic simulation uses geostatistics to generate a probability distribution for a spatial variable z, conditioned by available data including spatial autocorrelation in z and covariance with other spatial variables. Monte Carlo simulation is then used to sample this distribution and generate multiple alternative realizations (maps) of z that reflect the error

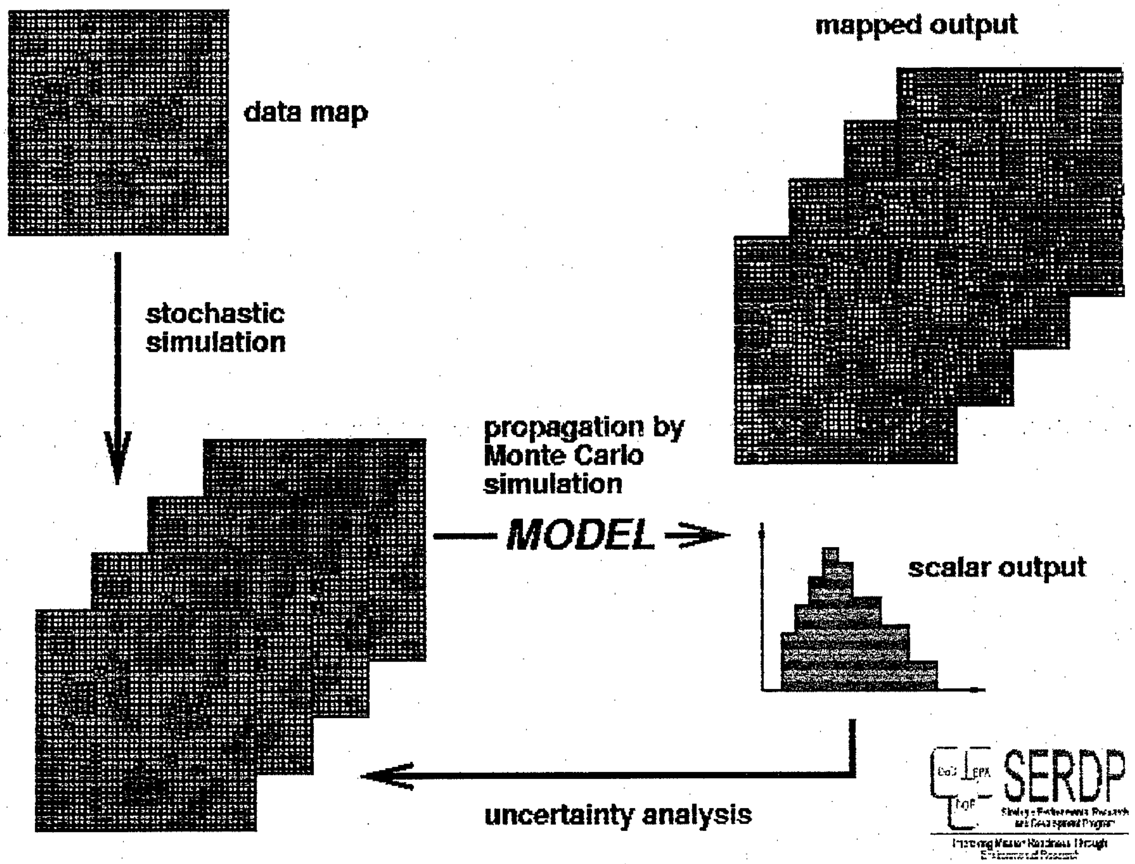

Figure 1: A framework for spatial error and uncertainty analysis of ecological models

and uncertainty in z. These maps are input to Monte Carlo simulation of the spatially structured ecological models utilizing that spatial data. Stochastic simulation is a widely recognized approach for addressing spatial uncertainty in geological and geographical applications (e.g., Geographical Information Systems). It is much less widely applied to ecological modeling, apparently because of lack of familiarity rather than any inappropriateness of the approach. We will adopt and adapt stochastic simulation to incorporate uncertainty in spatial data into ecological models. At this point, we have decided not to pursue alternative approaches to spatial uncertainty that involve fuzzy logic and fuzzy set theory. We feel these approaches are still too experimental and not well suited to DoD needs.

We have formulated a general approach or framework for the incorporation of uncertainty in spatial data into simulations with spatially explicit ecological models (Figure 1). The approach uses geostatistics and Monte Carlo simulation to propagate uncertainties in GIS data layers through the entire modeling process. This framework will be populated with the appropriate methods and tools. In 1998 we focussed on methods and tools for stochastic simulation to generate multiple realizations of one or more data maps (left of Figure 1).

As reported at the 1998 In-Progress Review, we have separated the treatment of (1) categorical spatial data (e.g., vegetation type) and (2) continuous data (e.g., population density). Different techniques of stochastic simulation are appropriate to each category. We focused on error in categorical data in Year 1 (1998), and we have identified *sequential indicator simulation* as the most appropriate method of stochastic simulation for categorical data in ecological models. We have elected to defer considerations of methods for continuous spatial data (e.g., Gaussian sequential simulation) until the second year of the project (FY 1999).

We have acquired and installed software libraries implementing methods for stochastic simulation of categorical spatial data. These include the GSLIB (Deutsch, 1998) and gstat Pebesma (1998) library. Both libraries are freely available and provide cost effective and portable tools. These libraries have been enhanced and supplemented with additional software developed by project participants B. L. Jackson and H. I. Jager. We have selected Arc/Info as the GIS component of our software implementation and as the user interface. While not yet fully implemented, in the future the stochastic simulation tools will be accessed by a user from the Arc/Info GIS environment.

We have applied these tools to a test data set of landcover type from Fort Knox, Kentucky. To summarize:

1. A 2km X 2km (100 X 100 pixel) subscene was extracted from a landcover GIS layer for Fort Knox, KY.

2. Ten percent of the submaps pixels were selected with a uniform, evenly spaced sampling scheme to generate a data set of 1000 sample points for input to the stochastic simulation routines. These routines utilize kriging algorithms that begin with a collection of sample points rather than a raster map.

3. The sample points were used as input to the categorical indicator simulation routines of the gstat library.

4. These routines generate multiple realizations (alternative versions) of the original input map, conditioned by the geostatistics (e.g., variogram) derived from the sample points.

Two realizations of the original submap are presented in Figure 2 as an example. Realizations A and B in Figure 2 are generated by slightly different methodologies. The coarser, larger scale spatial patterns and landscape structure of the original map are captured in the simulated maps. Table 1 illustrates how well the alternative methods are able to generate maps with landcover in proportions similar to that of the original or actual map. The landcover proportions in the generated maps generally agree with those of the original map, although as might be expected, the stochastic simulations seem to have more difficulty with the less common or rare landcover types. Similarly, certain landscape structures or patterns (e.g., narrow peninsula or corridors between patches) may not be well represented in the simulated maps. Increasing the density of the sampling points used as input to stochastic simulation might increase the chances of reproducing rare cover types or discrete fine scaled patterns. The potential gain has to be balanced, however, against the increased computational demand of a denser sampling with increased numbers of sample points.

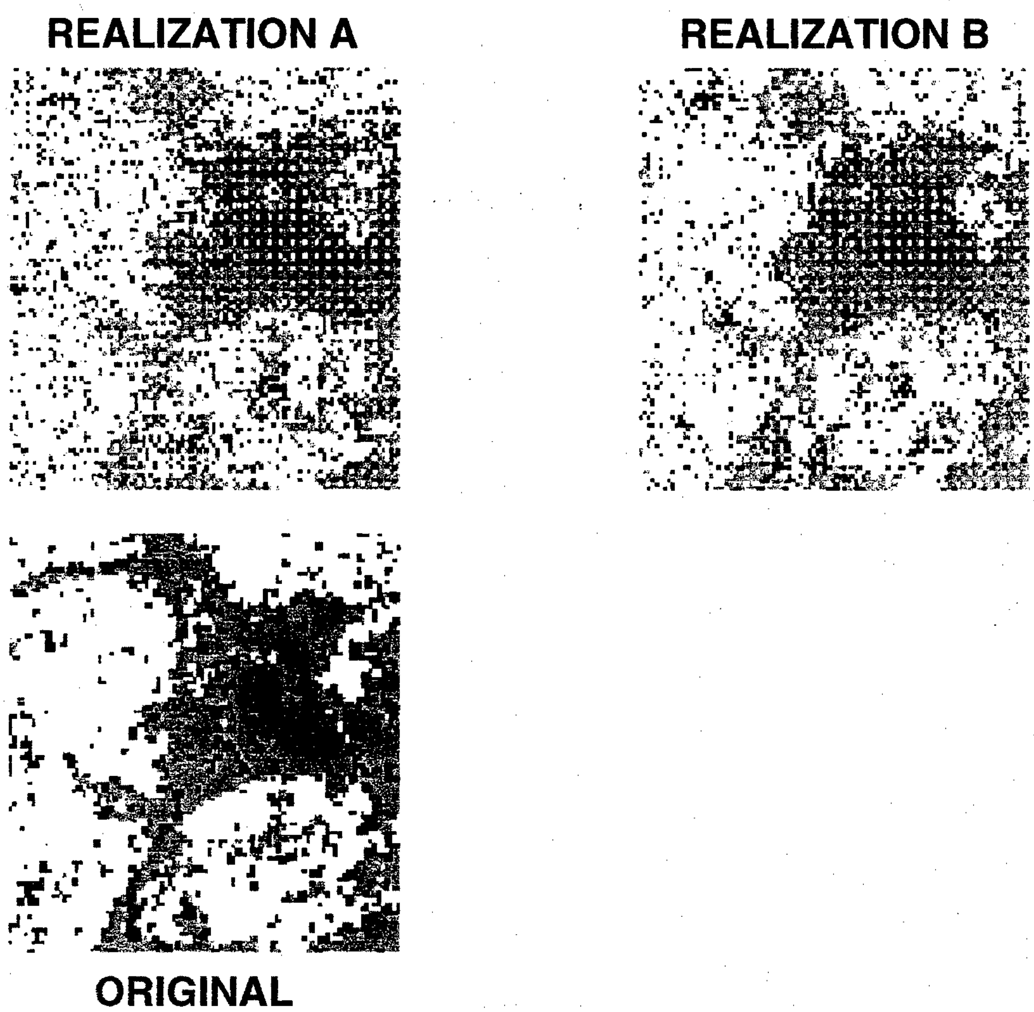

Figure 2: Alternative realizations by stochastic simulation of a landcover submap from Fort Knox, KY

Table 1: Landcover proportions in the original map and in realizations by stochastic simulation.

Land cover		Actual		loop100.ga			simlo100.ga		
Category	Description	Pixels	%	Pixels	%	Error	Pixels	%	Error
1	Deciduous forest	5,542	55.4%	4,612	52.0%	3.4%	4,841	54.6%	0.8%
2	Mixed forest	742	7.4%	785	8.9%	-1.4%	648	7.3%	0.1%
3	Evergreen forest	1	0.0%	7	0.1%	-0.1%	14	0.2%	-0.1%
4	Transitional	525	5.3%	562	6.3%	-1.1%	491	5.5%	-0.3%
5	Barren	2,339	23.4%	1,824	20.6%	2.8%	2,002	22.6%	0.8%
6	Urban	252	2.5%	191	2.2%	0.4%	197	2.2%	0.3%
7	Water		0.0%		0.0%	0.0%		0.0%	0.0%
8	Maintained grass	536	5.4%	721	8.1%	-2.8%	621	7.0%	-1.6%
9	Cropland	63	0.6%	160	1.8%	-1.2%	48	0.5%	0.1%
10	Lawn grass		0.0%		0.0%			0.0%	
Total		10,000	100.0%	8,862	100.0%		8,862	100.0%	

Our testing has revealed both the strengths and weaknesses of the software libraries. Challenges remain in the lack of efficiency in transferring data between software systems (e.g., Arc/Info and gstat) and with performing the simulation on the entire installation at Fort Knox. We are addressing the first challenge by investigating ways of minimizing and automating data transfer. We are addressing the challenge of slow execution times that result from the size of the spatial data sets by modifying, rewriting, and adding software where needed. We are making progress in solving these challenges.

Although our initial testing and development has been with data from Fort Knox, KY, we have selected Fort Hood, Texas as our case-study site. We will apply our spatially-structured avian population model (developed as part of CS-758) to populations of black-capped vireo and golden-cheeked warbler at Fort Hood, and we will complete an error and uncertainty analysis of the model focusing on the contribution of error and uncertainty in spatial data. The case study will guide the development of our general methods and tools for error and uncertainty analysis of spatial data and spatially explicit ecological models. The case study will also be used to test the appropriateness, generality, and portability of our approach and methods. Our familiarity with the avian population model will enhance the efficiency and productivity of the case study. The case study will familiarize us with spatial data at Fort Hood, facilitate coordination with the needs and requirements of the Army Training and Testing Area Carrying Capacity (ATTACC) methodology and of CS-1096, and will, as a side benefit, contribute to the investigation of black-capped vireo and golden-cheeked warbler concerns at Fort Hood. Our methods and tools (e.g., software) tested with the population model will be transferred to Dr. George Gertner and CS-1096 as part of Dr. Gertner's error budget approach for application to ATTACC at Fort Hood. The transfer will contribute to the population and ecological modeling components of the CS-1096 error budget and provide tools for application to ATTACC. ATTACC is presently unable to benefit from our methods because it is not a spatially explicit model. The methods and tools developed from our population model case study will be ready for future versions of ATTACC that are spatially explicit and require the methods we are developing. In addition, separating the case-study development of our methods and tools from the application of these tools to ATTACC will provide for a more powerful independent test and validation of our methods, one which will facilitate transfer and incorporation into the USACE Land Management System (LMS).

We have reviewed, evaluated, and assessed the approach and methodologies we identified for addressing spatial error and uncertainty in ecological models. We are satisfied that we have made the correct decisions and believe the challenges we have identified can be successfully overcome. We have elected to proceed with our project plan (a Go decision).

4.1 Technical Presentations

Dr. King and Mr. Ashwood presented a poster "Error and uncertainty analysis of spatially explicit ecological models" at the Partners in Environmental Technology 98, Technical Symposium and Workshop, December 1-3, 1998, Arlington, Virginia.

5 Issues

As noted above, our initial tests of the methods and software we are adopting and developing have revealed some issues that must be addressed. The first and most critical emerges when applying our tools to large spatial data sets characteristic of DoD installations. When applied to the entirety of Fort Knox for example we are faced with the challenge of large computer memory requirements and long execution times. We are addressing this issue with both hardware and software solutions. Our software tools have been recompiled and tested on a new SUN Ultra 5 workstation procured for this project. Execution times are significantly shorter with the new workstation. We continue to make refinements and revisions in our software that also provide improvements in execution time and efficiency. This issue must be resolved as we move to larger and larger installations like Fort Hood, Texas. The hardware solution has its limits, since the hardware demands of the software cannot exceed the capabilities of computational platforms available at most installations or of the systems that will be running LMS. Thus, we are focusing primarily on software solutions that can be implemented on the machines comparable to our Ultra 5 workstation found at most installations.

A related issue surrounds the ability of our methods to reproduce rare categorical types (e.g., rare habitat) and fine scale landscape structure (e.g., small patches and narrow corridors). These features are often of particular importance to ecological models describing resources of conservation concern. To be useful, our methods and tools must be able to recover these features of conservation concern with reasonable fidelity. Increasing the sampling density of the points used as input to the stochastic simulation will likely enhance our ability to represent these features, but this will increase the computational demands which are already challenging. Again we are investigating methodological and software solutions to this problem. For example, we are exploring a focussed stratified sampling scheme that has greater sampling density in rare types and surrounding finer-scaled structural features of particular interest.

6 Future Directions

Our objectives and milestones for 1999 are:

1. Acquire spatial data sets for case study at Fort Hood.
2. Parameterize our spatially-structured avian demographic model for black-capped vireo and golden-cheeked warbler at Fort Hood.
3. Identify and implement techniques for handling uncertainty in continuous and discrete quantitative spatial data.
4. Collect data to characterize uncertainty for case study.
5. Complete case study
6. Annual report

In addition to these milestones we will:

1. Continue our coordination with Dr. Gertner and CS-1096. We have planned a coordination meeting at Fort Hood in the March 1999 time frame.
2. Continue our investigation of solutions to the challenges posed by large spatial data sets and rare landscape features. The large area of Fort Hood and the large data sets that will be part of the case study provide an opportunity to address these issues.
3. We have submitted an abstract on our work from this project for a presentation at the 1999 Annual Meeting of the North American Chapter of the International Society for Ecological Modelling. We will also present portions of our work the international symposium "Predicting Plant, Animal, and Fungi Occurrences: Issues of Scale and Accuracy" in October 1999.

References

Burrough P.A. 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford University Press, New York.

Deutsch C.V. 1998. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, second edition.

Goodchild M.F. and Gopal S. (editors). 1989. *The Accuracy of Spatial Databases*. Taylor and Francis, London.

Isaaks E.H. and Srivastava R.M. 1989. *An Introduction to Applied Geostatistics*. Oxford University Press, New York.

Pebesma E.J. 1998. gstat 2.0g. http://www.geog.uu.nl/gstat/.

UCGIS. 1996. University consortium for geographic information science, reserarch prioritites for geographic information science. http://www.ucgis.org.

A Appendix: An Annotated Book Outline PERSPECTIVES ON UNCERTAINTY IN SPATIAL DATA FOR ECOLOGICAL ANALYSES

Editors: Carolyn Hunsaker, Michael Goodchild, Mark Friedl, and Ted Case
Table of Contents (7/16/98)

A.1 Part I INTRODUCTION AND OVERVIEW

Chapter 1 Authors: Michael Goodchild, et al.
Primary editor: Hunsaker; Secondary editors: Case and Friedl
Color figures: none
Status: draft complete, needs to be lengthened, editing beginning

This chapter sets the tone and content of the book by providing a framework for the role of spatial data, process models, and predictions in decision-making. The book is about the problems that uncertainty creates in this simple framework. It is virtually impossible to have certainty about the analysis process that leads to decisions, because in today's world decisions almost always involve multiple stakeholders with multiple viewpoints. Similarly, knowledge of ecological processes is never perfect, so uncertainty is present whenever outcomes are predicted. Finally, we can never have perfect knowledge of the boundary conditions, because the real world is far too complex to be fully measured, observed, or represented. This book covers all aspects of uncertainty, though its heaviest emphasis is on the third kind, uncertainty in the spatial data that provide the boundary conditions for ecological processes. We have chosen to place the emphasis there for several reasons: spatial data uncertainty is a rich area of research in several fields but this literature is relatively unknown in ecology, and the methods are not often applied; application in ecology will help to enrich the set of examples and motivating applications for the general study of spatial data uncertainty; and we believe the problems posed by spatial data uncertainty are among the most challenging, the most important for decision-makers, and may in some cases be the largest in magnitude.

A.2 Part II USE OF SPATIAL DATA IN ECOLOGICAL ANALYSIS

Chapter 2 Spatial Ecological Models
Authors: Fred Sklar and Carolyn Hunsaker
Primary editor: Case; Secondary editor: Goodchild
Color figures: 3 to 5
Status: draft in progress

This chapter reviews types of ecological models with emphasis on spatially explicit ones and the ways in which ecological models use spatial data. It will build upon previous reviews by Sklar and by Hunsaker. This chapter discusses the sources of uncertainty in spatial data and models: data collection, data processing, model structure, human intervention, and natural variability. It also suggests that despite a diversity of approaches, uncertainty will propagate through four components of model structure: inputs, initial conditions, calibration, and validation.

Chapter 3 Coastal Sage Scrub Case Study

Authors: Ted Case, Peter Stine, and Carolyn Hunsaker

Primary editor: Goodchild; Secondary editor: Friedl

Color figures: 1 or 2

Status: draft in progress

This chapter provides an example of a complex monitoring program for several sensitive terrestrial animal species and highlights the challenges ecologists are faced with when analyzing spatial data at different scales with the goal of determining overall ecological condition for several species. Data from this activity is used in one or more of the book chapters.

Chapter 4 Incorporating Uncertainties in Animal Location and Map Classification into Habitat Relationships Modeling

Authors: Kevin S. McKelvey and Barry R. Noon

Primary editor: Case; Secondary editor: Hunsaker

Color figures: none

Status: draft complete, editing beginning

Many of our understandings of animal habitat requirements are based on location data. We assume that those types of vegetation in which an organism can be consistently located represent important habitat. Deriving habitat associations from point-location data has traditionally been problematic. Errors in map classification combined with location error can lead to weak and biased interpretations of the data as shown by the simulations described in this chapter. The authors believe that their results show that the evaluation methods described here can lead to improved habitat relationship modeling, especially for organisms having strong associations with fairly uncommon vegetation types.

Chapter 5 Generic Issues Regarding Uncertainty in Spatial Data for Ecological Applications

Authors: Peter Stine and Carolyn Hunsaker

Primary editor: Case; Secondary editor: Goodchild

Color figures: none

Status: draft in progress

This chapter provides an interface between Part II where we describe how spatial uncertainty in ecological data can introduce errors in ecological analyses and Part III where we present methods that should provide ecologists with tools to better incorporate uncertainty into their research designs and analyses. This chapter includes discussion of organizational hierarchies and classifications ecologists use, examples of ecological issues where spatial pattern is relevant, availability of ancillary digital spatial data, and ways of describing and communicating uncertainty.

A.3 PART III METHODS

Cartography and Cognition

Chapter 6 Mapping Ecological Uncertainty

Author: Barbara Buttenfield

Primary editor: Goodchild; Secondary editor: Friedl

Color figures: none

Status: draft in progress

This chapter briefly describes the basic information and classical ways in which cartographers represent spatial entities. In general uncertainty has not been incorporated into cartographic techniques. The author presents some approaches that cartographers might use in the future to represent uncertainty in their products.

Chapter 7 A Cognitive View of Spatial Uncertainty

Authors: Geoffrey Edwards and Marie-Josee Fortin

Primary editor: Goodchild; Secondary editor: Friedl

Color figures: none

Status: Draft complete, editing beginning

The conceptualization of a category is rooted in human cognition and perception, although it may be socially or culturally systematized or incorporated within a mathematical framework. A representation of space in terms of idealized geometrical objects such as polygons, lines, and points is merely one, rather simplified approach. Standard models and representations are poorly suited to the representation of more complex spatial distributions such as ecological systems, characterized by continuous and dynamically changing interactions at many scales simultaneously. The authors explore some of the links between spatial uncertainty and human perception and cognition. The different processes by which data are collected and transformed is examined. Different forms of spatial uncertainty which have a cognitive component and are likely to be pertinent to ecological data are highlighted, and currently emerging techniques for handling such uncertainties are surveyed. The authors examine new, longer term research which seeks to formalize cognitive and perceptual representations and hence to provide new means of addressing spatial uncertainty and the role it plays in constraining data used for decision-support.

Spatial Statistics

Chapter 8 Spatial Analyses of Ecological Data

Authors: Marie-Josee Fortin and Geoffrey Edwards

Primary editor: Goodchild; Secondary editor: Friedl

Color figures: none

Status: draft complete, editing beginning

This chapter addresses how the intrinsic spatial dependencies of spatial data affect landscape pattern quantification, especially boundary detection, according to data sources: field (sampling), aerial photograph, and remote sensing. A quick overview of the inherent properties of spatial data is presented (spatial dependence-autocorrelation, grain, extent, multiple scales, etc.). Advantages and disadvantages of boundary delineation methods according to the sources of the data (field, aerial photograph, remote sensing) are illustrated.

Chapter 9 Geostatistical Models of Uncertainty for Spatial Data

Author: Phaedon C. Kyriakidis

Primary editor: Friedl; Secondary editor: Goodchild

Color figures: none necessary but would improve understanding on some

Status: draft complete, editing beginning

Ecologists and environmental scientists are frequently faced with the task of predicting attribute levels, such as the density of a population or the concentration of a pollutant at unsampled locations. This chapter discusses the use of geostatistics for data integration and stochastic simulator-performance assessment. Ground-based measurements and remotely sensed information are integrated via geostatistical algorithms to generate maximally-constrained simulated realizations, which are input in an environmental process simulator. The resulting multiple simulated outcomes of the unknown process response allow assessing the impact of the uncertainty regarding the spatial model on the uncertainty regarding the true response value. The random function model and related aspects of spatial dependence and spatial correlation are introduced. The kriging paradigm for updating prior uncertainty models into data-conditioned models of local uncertainty is presented. Various integration avenues are introduced for incorporating secondary information, such as remotely sensed imagery, in order to further constrain local conditional uncertainty models. The sequential simulation paradigm for assessment of joint spatial uncertainty is described along with some conceptual and methodological issues.

Chapter 10 Spatial Linear Models in Ecology

Authors: Jay M. Ver Hoef and Noel Cressie

Primary editor: Goodchild; Secondary editor: Friedl

Color figures: none

Status: draft complete, editing beginning

This chapter demonstrates the usefulness of the spatial linear model for making inferences from ecological data. The spatial linear model is at the heart of many spatial methods, from optimal prediction to designed experiments. Robust methods are needed to deal with ecological data, and there are two approaches. One is to use robust estimation methods and the other is to assume robust models. Statistical models like linear regression posit simple relationships among the variables that may be unrealistic, but have the virtue that the uncertainty of the estimated parameters in the model can be easily quantified. This chapter shows that the spatial linear model is robust. Data from the case study presented in chapter 3 are used here.

Chapter 11 Characterizing Uncertainty in Digital Elevation Models

Author: Ashton Shortridge

Primary editor: Friedl; Secondary editor: Goodchild

Color figures: none

Status: draft complete, editing beginning

Topography plays an important role in many environmental processes. Discrepancies exist between digital elevation models (DEMs) and the real-world surfaces they represent. Using an uncertainty model, a researcher can propagate DEM uncertainty through the analysis to identify its impact upon the results of the application. This is accomplished by producing, via Monte Carlo simulation, a set of equiprobable realizations of the DEM. This chapter provides a through discussion of DEMs and uncertainty, as well as indicating general approaches to modeling uncertainty in data for continuous phenomena. Some examples are presented to illustrate these modeling approaches.

Chapter 12 Uncertainty of Multinominal Spatial Data

Author: Charles R. Ehlschlaeger

Primary editor: Friedl; Secondary editor: Goodchild

Color figures: none

Status: rough draft complete, editing beginning

Multinominal maps are defined as polygons with homogenous characteristics such as vegetation cover or soil classes. There are two conceptual representations of multinominal maps: field based which is more closely related to raster data structures and object based which is more closely identified with vector data. This chapter takes a field-based approach using a raster GISystem. It formulates a methodology, distributive-parametric stochastic simulation, that makes it possible to use lower-quality data and return a distribution of model results. Basically information from a sample of high-quality maps within an area having continuous data only of coarser quality is used to model representations at the higher quality for the entire area. This distribution of multinominal maps can then be used in ecological models requiring spatial data and thus provide a distribution or probability estimate from the ecological model. This chapter uses an example based on the case study presented in chapter 3, a habitat model for the California gnat catcher. Given a good understanding of the ecological process model, spatial data uncertainty modeling is computationally viable, theoretically complete, and affordable by most data development budgets.

Remote Sensing

Chapter 13 An Overview of Uncertainty in Remote Sensing for Ecological Applications

Authors: M.A. Friedl, K. McGwire, and F.W. Davis

Primary editor: Goodchild; Secondary editor: Hunsaker

Color figures: unknown at this time

Status: outline, draft nearing completion

Remote sensing has become a widely used tool for a variety of ecological applications including multi-species habitat mapping, land cover and land use change monitoring, estimation of carbon assimilation rates and net primary production. The experience of many ecologists who have attempted to use remote sensing over the past two or three decades suggests that a danger exists for misuse and over-optimism regarding the true utility of remote sensing data sources for problems in ecology. The purpose of this chapter is to provide an overview of the limitations inherent to remote sensing with the goal of providing ecologists guidance regarding the appropriate use of this rather complex source of data. To this end the chapter is composed of two main elements: the nature of remote sensing and key concepts that pervade the use of remote sensing for ecological applications. We focus on the interaction between data spatial resolution and the scale of ground scene properties, as this relationship determines the ultimate utility of a remote sensing data set for a specific application. The chapter concludes by discussing future sensor systems and their role in ecological science.

Chapter 14 Remote Sensing Classification of Forest Covertype and Estimation of Stand Leaf Area Index for Modeling Net Primary Production

Author: S.E. Franklin

Primary editor: Friedl; Secondary editors: Goodchild, Hunsaker and Case

Color figures: 6 plates

Status: draft complete, editing beginning

To estimate actual net primary production (NPP) timely observations on covertype and leaf area index (LAI) are required for use in calculations of nutrient cycling and to estimate the amount of woody biomass per site. This estimate, in turn, is considered in the model estimates of stand respiration and PAR conversion efficiency. The actual LAI for a given covertype is a decisive factor in estimates of photosynthesis and litter decomposition or turnover rates. This chapter shows how satellite remotely-sensed estimates of covertype and LAI are superior to GIS-based covertype labels, often derived through aerial photointerpretation, and assumptions of maximum LAI, often based on climate and other environmental constraints. In the example described for the mixedwood Acadian Forest Region of New Brunswick's Fundy Model Forest, estimates of NPP varied by as much as 25difference between assumed (climate-driven) LAI and actual LAI measured by remote sensing, and by as much as 50assumed GIS polygon homogeneity and actual polygon heterogeneity measured by remote sensing.

Chapter 15 Spatially Variable Thematic Accuracy: Beyond the Confusion Matrix

Authors: Kenneth McGwire and Peter Fisher

Primary editor: Friedl; Secondary editor: Goodchild

Color figures: none

Status: rough draft complete, beginning editing

An essential aspect of the increasing sophistication of ecological models is the use of methods with spatially explicit inputs and outputs. Thus, traditional challenges in documenting the uncertainty of model parameters are expanding to include uncertainty in spatially disaggregated data inputs. Fortunately, it is becoming more common for spatial data inputs, such as those derived from remote sensing data, to have their overall map accuracy documented. However, for complex, spatially explicit models the distribution of error over a geographic domain becomes very significant in determining the validity of model outputs. The most effective way to assess these types of models is generally through a Monte Carlo analysis, since an analytical assessment of the impacts of spatially varying errors in multiple data inputs may not be possible. However, a problem for implementing such Monte Carlo methods arises since the commonly accepted method for assessing the accuracy of thematic maps, the confusion matrix, is entirely devoid of spatial context. This paper addresses shortfalls in attempting to apply the confusion matrix developed for an extensive thematic map to a specific subregion, indices for documenting spatial pattern, and how the traditional error matrix and indicators of spatial pattern can be combined to realistically simulate the landscape in a Monte Carlo analysis.

Chapter 16 Modeling Spatial Variation of Classification Accuracy Under Fuzzy Logic

Author: A-Xing Zhu

Primary editor: Friedl; Secondary editor: Goodchild

Color figures: none

Status: draft complete, editing beginning

This chapter presents an approach to modeling spatial variation of classification accuracy in categorical maps. A similarity model based on fuzzy logic for representing spatial variation of geographic entities is recommended for modeling the spatial variation of classification errors. Under this model, commission errors are approximated by membership exaggeration (exaggeration uncertainty) and omission errors are estimated by membership ignorance (ignorance uncertainty). Two case studies in Montana were conducted to illustrate this approach: soil mapping in the Lubrecht Experimental Forest and surface cover mapping at Glacier National Park. The results of these two case studies indicate that areas of transition (mixed-grade and intergrade) have high ignorance uncertainty and moderate exaggeration uncertainty values. Areas covered with undefined classes (extra-grade) have high exaggeration uncertainty while areas of single and typical classes (typical-grade) have both low ignorance and exaggeration uncertainty.

Chapter 17 Set Theoretic Approaches to Uncertainty in Spatial Information

Author: Peter Fisher

Primary editor: Goodchild; Secondary editor: Friedl

Color figures: ?

Status: draft in progress

This chapter examines the set theoretical basis for uncertainty in spatial information of natural resources, and presents the diversity of set theories, which are now available with which to explore and understand uncertainty. The discussion emphasizes the distinction between Boolean and Fuzzy sets in spatial information. Speculations are made about the status of Rough Sets as an alternative to Boolean sets, with some emphasis on the difference between Rough and Fuzzy sets. It is suggested that while error in the assignment of objects to Boolean Sets is widely identified, there is a comparable error associated with estimated memberships of objects in Fuzzy sets. This is one aspect of higher order uncertainty, which has been identified as a fundamental property of any vague set, such as a Fuzzy set.

Data Integration and Decision Making

Chapter 18 Roles of Meta-Information in Uncertainty Management

Author: Kate Beard

Primary editor: Goodchild; Secondary editor: Friedl and Hunsaker

Color figures: none

Status: Draft complete, editing beginning

Meta-information is information that describes information, i.e., information that makes data useful. The significance of meta-information is most apparent in situations where datasets are widely distributed as in digital libraries or under any circumstances where the data user is not the data collector and is thus less likely to be familiar with characteristics and idiosyncrasies of the data. Meta-information serves the function of providing information to find information as in a distributed on-line environment and to effectively use such information once it has been discovered. Uncertainty is not a characteristic of data or information but rather a state of knowledge regarding the data. Meta-information has the potential to directly reduce some amount of uncertainty. Uncertainty is inherent in ecological analysis and synthesis. Sources of uncertainty include non-deterministic ecological processes; conceptual vagueness or ambiguity; sampling; data collection errors which lead to incomplete, incorrect, or imprecise observations; use of surrogate variables; representational constraints; and data processing. This chapter reviews sources of uncertainty and identifies the roles of meta-information in mitigating these sources. It also reviews current concepts of meta-information and discusses how these can be extended to improve uncertainty management. Uncertainty management refers to improving comprehension and understanding of imperfections in data and information.

Chapter 19 Making Decisions Under Uncertainty Using GIS

Author: Ron Eastman
Primary editor: Goodchild; Secondary editors: Friedl and Hunsaker
Color figures: possibly
Status: draft in progress
This chapter provides examples of how spatial data is used in geographic information systems (GIS) for making decisions about environmental issues and explores the importance of uncertainty in such analyses.

A.4 Part IV Epilog

Chapter 20 Authors: Hunsaker, et al.
Primary editor: Goodchild; Secondary editors: Case and Friedl
Color figures: none
Status: to be written once all chapters are available
This chapter will briefly review the primary points discussed and conclusions drawn in each chapter and provide a synthesis of the state of our knowledge regarding uncertainty in spatial data for ecological analyses. It will also reiterate critical connections that need to be made by ecologists between spatial data acquisition, representation, manipulation, and use in analyses. Thus we hope to bring together in a new way the needs of ecologists with the current tools from research in cartography, cognition, spatial statistics remote sensing, and computer sciences with regard to quantification and description of uncertainty in spatial data.